AFML-TR-73-94

AD 762546

THE EKS-SQUARE TEST OF GOODNESS OF FIT– AN IMPROVEMENT OF THE CHI-SQUARE TEST

WALODDI WEIBULL
1012 LAUSANNE
SWITZERLAND

TECHNICAL REPORT AFML-TR-73-94

MAY 1973

AIR FORCE MATERIALS LABORATORY
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

AIR FORCE/56780/25 June 1973 — 250

THE EKS-SQUARE TEST OF GOODNESS OF FIT–AN IMPROVEMENT OF THE CHI-SQUARE TEST

WALODDI WEIBULL
1012 LAUSANNE
SWITZERLAND

Approved for public release; distribution unlimited.

FOREWORD

The research work reported herein was conducted by Prof. Dr. Waloddi Weibull, Chemin Fontanettaz 15, 1012 Lausanne, Switzerland under USAF Contract No. F44620-72-C-0028. This contract, which was initiated under Project No. 7351, "Metallic Materials", Task No. 735106, "Behavior of Metals", was administered by the European Office, Office of Aerospace Research. The work was monitored by the Metals and Ceramics Division, Air Force Materials Laboratory, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, under the direction of Mr. W. J. Trapp, AFML/LL.

This report covers work conducted during the period 1 February 1971 to 25 February 1972. The manuscript was submitted for publication by the author in March 1972.

This technical report has been reviewed and is approved.

W. J. Trapp

W. J. TRAPP
Act Asst for Reliability
Metals and Ceramics Division
Air Force Materials Laboratory

ABSTRACT

When applying the classical Chi-square test of goodness of fit, it is always assumed that the test statistic is χ^2-distributed. Since this is true only for very large samples, some restrictions on the class frequencies have to be introduced. It is generally accepted that none of the expected frequencies should be less than ten, which makes this test useless for small and moderate samples.

In order to eliminate these - from a practical viewpoint severe - restrictions, it is proposed to use the exact sampling distribution instead of the limiting χ^2-distribution. When doing so, the test will be called the Eks-square test.

Programs have been written for computing these distributions and the improvements attained have been stated.

The possibilities of using the modified test statistic as a location, scale, and shape operator have been examined and illustrated by numerical examples. Several tables have been prepared.

TABLE OF CONTENTS

LIST OF TABLES

I

INTRODUCTION

The most generally used test of goodness of fit for the last seventy years, called the Chi-square test, was introduced by K. Pearson [1]. Its test statistic X^2 is defined by

$$X^2 = \sum_1^r [(v_i - N.p_i)^2/N.p_i] \tag{1}$$

where

r = a finite number of parts (classes) without common points into which the space of the variable has been divided,

p_i = the corresponding values of the given probability function ($\Sigma p_i = 1$),

v_i = the observed class frequencies in the sample of size N ($\Sigma v_i = N$).

Introducing the expected class frequencies

$$v_{oi} = N.p_i \tag{2}$$

we have

$$X^2 = \sum_1^r [(v_i - v_{oi})^2 v_{oi}] = \sum_1^r [v_i^2/v_{oi}) - N \tag{3}$$

In particular, if all the class frequencies are equal to v_o, we have

$$X^2 = \sum_1^r v_i^2/v_o - N \tag{4}$$

where

$$r = N/v_o \tag{5}$$

Introducing (5) into (4) results in

$$X^2_{r/N} = r\sum_{1}^{r} v_i^{\,2}/N - N \tag{6}$$

A remarkable property of the statistic X^2 is due to the fact that the standardized variable

$$w = (v - N.p)/\sqrt{N.p.q} = (v - v_o)/\sqrt{v_o(N - v_o)/N} \tag{7}$$

is asymptotically normal $(0,1)$ on the condition that p remains constant when $N \longrightarrow \infty$. Hence, on certain conditions the random variable

$$w = (v - v_o)/\sqrt{v_o} \tag{8}$$

tends to the random variable ξ , which is normal $(0,1)$.

Thus,

$$X^2 \longrightarrow \sum_{1}^{r} \xi^2 \tag{9}$$

that is, X^2 is, in the limit, distributed in a χ^2-distribution with $r-1$ degrees of freedom (d.fr.).

The fact that it is always assumed that X^2 is χ^2-distributed, makes it necessary to introduce some restrictions. The following conditions are generally accepted (Cf.Cramér [2], p.420): When the χ^2-test is applied in practice, and all the expected frequencies $N.p_i$ are ≥ 10, the limiting χ^2-distribution can be used with an approximation sufficient for ordinary purposes. If some of the $N.p_i$ are < 10 it is advisable to pool the smaller classes, so that every class contains at least 10 expected observations, before the test is applied. When the observations are so few that this cannot be done, the χ^2 tables should not be used, but some information may still be drawn from the values of the mean and the variance of

The condition $v_o = N.p_i \geq 10$ restricts, in fact, the use of this test to quite large samples. Considering that there must be at least one degree of freedom and that the number of degrees of freedom is reduced by one unit for each parameter

estimated from the sample, it follows that for a two-parametric hypothetical distribution with unknown parameters, say, for the normal distribution, the sample size should not be less than $N = 40$, and for a distribution with three unknown parameters not less than $N = 50$.

Even these, from a practical view-point rather severe restrictions, are in some cases not sufficient as will be demonstrated in the following studies of the exact distributions of the statistics w and X^2.

It is evident that, if the exact distribution of X^2 is used instead of the limiting χ^2-distribution, then there will be no need of restrictions with regard to v_o. In view of the fact that any pooling of classes implies a reduction of the amount of information provided by the sample, as will be proved in the following, it seems, in some cases, desirable to use as many and as small classes as possible. Thus, if all sample values are known, then $v_o = 1$ will sometimes be the preferable value, and, if the sample is presented as a table with given class limits and corresponding observed frequencies v_i, then no pooling of the classes would be undertaken.

In order to distinguish between the two alternatives of using either the limiting χ^2-distribution or the exact distribution of the test statistic X^2, these two types of tests will be called the Chi-square and the Eks-square test of goodness of fit. The X^2-test will, in general, operate with much smaller values of v_o, even $v_o = 1$, than will the χ^2-test.

Since the use of the χ^2-distribution as an approximation to that of X^2 is based on the assumption that the random variable $w = (v - v_o)/\sqrt{v_o}$ is normally (0,1) distributed, it may be of interest to state its deviations from normality. To this purpose, some exact distributions of w have been deduced and examined as indicated in the following section.

II
THE EXACT DISTRIBUTION OF $W = (V - V_0)/\sqrt{V_0}$

Let the probability that one single observation falls within the i:th of r classes be p_i $(i = 1,2,\ldots,r; \Sigma p_i = 1)$, then the probability that N independent observations are distributed in such a way that there are v_i observations within the i:th class, becomes

$$p = \frac{N!}{v_1! \, v_2! \ldots v_r!} \cdot p_1^{v_1} \cdot p_2^{v_2} \cdot \quad p_r^{v_r} \tag{10}$$

For the special case that we have only two classes with the probabilities p and $1-p$, then the probability that v observations fall within the first class and $N-v$ within the second class is

$$p(v) = \frac{N!}{v!(N-v)!} \cdot p^v (1-p)^{N-v} \tag{11}$$

Introducing

$$p = v_o/N \tag{12}$$

we have

$$p(v,v_o) = \frac{N!}{N^N} \cdot \frac{v_o^{\,v}}{v!} \cdot \frac{(N-v_o)^{N-v}}{(N-v)!} \tag{13}$$

For large values of N the right-hand member of equ.(13) is a quotient of very large integers, which makes it difficult to compute the value of $p(v,v_o)$ with sufficient precision.

For this reason the following three recurrence formulas have been deduced.

a) v_o = constant

$$p(v+1,v_o) = p(v,v_o) \cdot \frac{v_o}{v+1} \cdot \frac{N-v}{N-v_o} \tag{14}$$

b) v = constant

$$p(v,v_o+1) = p(v,v_o) \cdot \frac{(v_o+1)^v}{v_o^{\,v_o}} \cdot \left(1 - \frac{1}{N-v_o}\right)^{N-v} \tag{15}$$

c) $v = v_o$ –

$$p(v_o+1,v_o+1) = p(v_o,v_o) \cdot \frac{(v_o+1)^{v_o}}{v_o^{\,v_o}} \cdot \left(1 - \frac{1}{N-v_o}\right)^{N-v_o} \tag{16}$$

From (14) it is easily found that $p(v_o,v_o)$ is the supremum of each v_o-column and from (15) that it is the supremum also of each v-row. Further, it follows from (16) that $p(v_o,v_o)$ is monotonously decreasing with v_o.

Based on the preceding formulas, Program 5/71 has been written. Several tables have been prepared for sample sizes up to $N = 10{,}000$. Some of the results are presented in Table 1 and Table 2. Here the function $F(x)$ is equal to the probability that $w \leq x$. The following two important formulas are verified by the tables.

The expected value of w

$$E(w) = 0 \tag{17}$$

The expected value of w^2

$$E(w^2) = (N - v_o)/N = (r - 1)/r \tag{18}$$

In Table 1 the exact distribution $F(w)$ and the normal distribution $\phi(w)$ are listed for $v_o = 1$ and several sample sizes N up to $N = 10{,}000$. There is a principal difference between these two distributions in so far as the former is a discrete and the latter a continuous function. The sample size N has very little effect on $F(w)$, for $N \geq 50$ practically none at all.

In Table 2 the functions $F(w)$ and $\phi(w)$ are listed for $N = 1{,}000$ and $v_o = 1, 4,$ and 16. Also for the last value of v_o there is a definite difference between the two distributions, as conspicuously demonstrated in Fig.1.

This results motivate a closer examination of the X^2-distributions and their deviations from normality.

We then have to distinguish between two cases:

a) the hypothetical distribution is completely specified

b) certain parameters of the hypothetical function are estimated from the sample.

III

THE HYPOTHETICAL DISTRIBUTION IS COMPLETELY SATISFIED

Let the hypothetical distribution be specified by the cumulative distribution function (Cdf)

$$F[(x-\mu)/\beta] \tag{19}$$

with all parameters known.

If now our sample is presented as a table with fixed class limits x_c and observed class frequencies v_i, then the expected class frequencies are

$$v_{oc} = \left\{F[(x_c - \mu)/\beta] - F[x_{c-1} - \mu)/\beta]\right\}N \tag{20}$$

and we have merely to calculate the test value X^2 by introducing the two sets v_i and v_{oc} into the formula (3).

In order to decide whether this test value corresponds to an acceptable goodness of fit, the sampling distribution of X^2, specified by the set v_{oc}, has to be known. This distribution can be determinated by means of a Monte-Carlo procedure according to Program 21/71. It provides the probability of obtaining an value X^2 which is larger than the test value. This program will be described in details in the following. It should only be mentioned here that generating 10,000 random samples from the hypothetical population, computing for each sample a random value X^2 and classifying them takes a computing time less than 20 scs. for a sample of size $N = 10$ and less than 40 scs. for $N = 20$.

If all the elements x_i of the sample are known, then we can freely choose the expected frequencies v_{oc}. The most convenient choice will be to let all expected class frequencies be the same, which makes r classes, each having the expected frequency $v_o = N/r$. The value of $X^2_{r/N}$ is then given by equ.(6).

This particular case, will be more closely examined.

3.1 General properties of the statistic $X^2_{r/N}$ –

From the definition (6) it immediately follows that

$$\inf X^2_{r/N} = 0 \quad ; \quad \sup X^2_{r/N} = N(r-1) \tag{21}$$

Since the infimum of X^2 corresponds to $v_i = N/r$ and $p_i = 1/r$, we have from (10)

$$\mathrm{Prob}(X^2_{r/N} = 0) = N!/r^N \cdot (N/r!)^r \tag{22}$$

which for $v_o = 1$, $r = N$ becomes

$$\mathrm{Prob}(X^2_{N/N} = 0) = N!/N^N \tag{23}$$

Since the supremum of X^2 corresponds to the case that anyone of the r classes contains N observations we have from (10)

$$\mathrm{Prob}[X^2_{r/N} = N(r-1)] = r^{-(N-1)} \tag{24}$$

which for $v_o = 1$, that is $r = N$, becomes

$$\mathrm{Prob}[X^2_{N/N} = N(N-1)] = N^{-(N-1)} \tag{25}$$

It is a remarkable fact that the expected value

$$E(X^2_{r/N}) = r - 1 = E(\chi^2) \tag{26}$$

and the variance

$$\mathrm{Var}(X^2_{r/N}) = 2(r-1)(N-1)/N = \mathrm{Var}(\chi^2) \tag{27}$$

that is, in spite of quite different distributions, the expected values and variances of $X^2_{r/N}$ and χ^2 are identical.

These two formulas have not been theoretically deduced, but they are exactly verified by use of the theoretically deduced distributions of $X^2_{r/N}$ and closely approximated by use of the Monte-Carlo determined distributions, as demonstrated below. Equ.(26) is, however, a corollary of equ.(18).

3.2 Some exact distributions of $X^2_{r/N}$ –

From equ.(6) it follows that for given r, N the quantity $X^2_{r/N}$ is uniquely determined by the set of observed frequencies

v_i. In the actual case, the probability of obtaining this set is, introducing $p_i = 1/r$ in equ.(10)

$$p = N!/r^N . v_1! \, v_2! \, .. \, v_r! \tag{28}$$

Since the value of $X^2_{r/N}$ is independent of any permutation of the frequencies, we have to multiply the probability p by the number of different permutations of v_i.

For small values of r this is a feasible task. In this way, the distributions of $X^2_{2/10}$, $X^2_{2/20}$ and $X^2_{3/9}$ have been calculated. The results are presented in Table 3.

For large r it is more convenient to determine the distributions by use of a Monte-Carlo studies in the following way.

Let the hypothetical distribution be

$$P_i = r_i = F[(x_i - \mu)/\beta] \tag{29}$$

Putting $P_i = r_i$, where r_i is a random variable uniformly distributed on the interval (0,1) is made in order to emphasize the following, most important property of any continuous distribution function $F(x)$. (Cf.Wilks[4],p.13): If X is a random variable having a continuous cdf $F(x)$ then $F(X)$ is a random variable such that

$$\text{Prob}\,[F(X) = p] = p \tag{30}$$

Hence, the random variables P_i are independently and uniformly distributed on the interval (0,1), just as are r_i.

Inverting equ.(29) we have

$$x_i = \beta . F^{-1}(r_i) + \mu \tag{31}$$

If now one of the classes, into which the space of the variable x has been divided, has the limits x_a and x_b and if x_i falls within this class, that is, if

$$x_a < x_i < x_b \tag{32}$$

then

$$r_a = F[(x_a - \mu)/\beta] < r_i = F[(x_i - \mu)/\beta] < r_b = F[(x_b - \mu)/\beta] \qquad (33)$$

because the function F is nondecreasing. From (32) and (33) it can be concluded that the probability of x_i falling within the interval (x_a, x_b) is equal to that of r_i falling within the interval (r_a, r_b).

Thus, if the values of the given sample have been arranged for tabulation purposes into an arbitrary number of classes with the limits $-\infty, x_1, x_2, \ldots, \infty$, then the corresponding number $0, r_1, r_2, \ldots\ldots 1$ are computed by use of (29), from which it also can be concluded that the expected frequency v_{oi} of the i:th class is

$$v_{oi} = N(r_i - r_{i-1}) \qquad (34)$$

From the preceding it follows that, instead of generating a random sample of size N from the population, defined by the cdf $F(x)$, and counting the number of elements x_i falling within each of the classes in the space of x, identical result is obtained by taking out at random a set of N random numbers r_i and counting the number of them which are falling within each of the classes in the space of r.

It is evident that, if the expected frequency of each class is equal to $v_o = N/r$, then the M.-C.-procedure is independent of the hypothetical distribution. The only way it enters into the problem consists in the calculation of the class limits by use of equ.(31).

The M.-C.-procedure is performed by Program 21/71. For any given sample size N, it generates a large number N_{part}, say, 10,000, of random samples r_i, computes from each of them $X^2_{r/N}$ for a selected number of v_o, for instance, for $N = 10$, $v_o = 1, 2, 5$, and for $N = 20$, $v_o = 1, 2, 5, 10$. Because the same random samples are used for the different values v_o much computing time is saved, since the frequencies corresponding to $v_o = 1$ are obtained by pooling the N classes for $v_o = 1$. Also the means and variances of $X^2_{r/N}$ are computed and written down.

The computing times were

59.0 sec for $N = 10$; $r = 2, 5, 10$; $N_{part} = 30{,}000$

48.6 sec for $N = 20$; $r = 2, 4, 10, 20$; $N_{part} = 10{,}000$

The distributions for $N = 10$ are presented in Table 4 and a comparison between $X^2_{2/10}$ and χ^2 for one degree of freedom in Fig.2.

3.3 Errors in the level of significance due to the assumption that $X^2_{r/N}$ is χ^2-distributed

As long as the X^2-distribution is assumed to be distributed in a χ^2-distribution with $r-1$ degrees of freedom, (d.fr.) the value X^2_p corresponding to a given level of significance p is assumed to be equal to the well known and tabulated values χ^2_p. Some of these values are listed below.

Level of significance p %	Values of χ^2_p			
	1 d.fr.	2 d.fr.	9 d.fr.	19 d.fr.
5	3.841	5.991	16.919	30.144
2	5.412	7.824	19.679	33.687
1	6.635	9.210	21.666	36.191
0.1	10.827	13.815	27.877	43.820

The p percent value χ^2_p of χ^2 for r d.fr. is a value such that the probability that an observed value of χ^2 exceeds χ_p is

$$\mathrm{Prob}(\chi^2 > \chi^2_p) = p/100$$

The error committed by using these values instead of the exact X^2_p can be stated by reading from the exact step functions $Q(x) = \mathrm{Prob}(X^2_{r/N} > x)$ the probabilities corresponding to the assumed values of p. Some errors are presented below.

Assumed level %	Exact levels of significance of				
	$X^2_{2/10}$ %	$X^2_{2/20}$ %	$X^2_{3/9}$ %	$X^2_{10/10}$ %	$X^2_{20/20}$ %
5	2.148	4.139	5.045	3.86	3.42
2	2.148	1.182	2.484	2.27	2.15
1	0.195	1.182	0.290	1.26	0.96
0.1	0.000	0.040	0.015	0.16	0.25

For example, the assumed level of significance 5% is actually 2.148% if the test statistic $X^2_{2/10}$ is used and X^2_p is taken = 3.841.

It may also be noted that $X^2_{2/20}$, as having $v_o = 10$, satisfies the accepted rule $v_o \geqq 10$. Nevertheless, there is a definite difference between the assumed and the true level of significance.

3.4 The statistic $X^2_{r/N}$ as a location or a scale operator

If the hypothetical distribution and one of the two parameters, location μ and scale β, are specified, then the other parameter can be selected using $X^2_{r/N}$ as a test operator.

The decision power of this operator has been determined by means of Program 22/71.

The principle on which this program is based consists in specifying completely two distributions, including the function, the location and the scale parameter. From the first distribution the class limits for an arbitrary number of equivalent classes r are computed. Then a large number of random samples, say 10,000, belonging to the population, specified by the second distribution, are generated. For each sample the value of $X^2_{r/N}$ is computed and the frequency distribution of these 10,000 values are determined. If the two distributions are identical, this program produces the same result as Program 21/71.

The following five alternatives have been run for $N = 10$; $r = 2, 5, 10$.

Distribution	Function	Alt.1 μ	Alt.1 β	Alt.2 μ	Alt.2 β	Alt.3 μ	Alt.3 β	Alt.4 μ	Alt.4 β	Alt.5 μ	Alt.5 β
1	Normal	0	1	0	1	0	1	0	1	0	1
2	Normal	1	1	2	1	3	1	0	0.5	0	2

By comparing these fifteen frequency distributions with those corresponding to Distribution nr 1, computed by use of Program 21/71, the following decision powers are obtained.

Statistic	$\mu = 0$ vs. 1	0 vs. 2	0 vs. 3	$\beta = 1$ vs. 0.5	1 vs. 2
$\chi^2_{2/10}$	$\boxed{68.5}$	96.1	99.94	0	0
$\chi^2_{5/10}$	66.0	$\boxed{97.8}$	99.95	$\boxed{48.3}$	33.5
$\chi^2_{10/10}$	58.7	97.5	$\boxed{99.98}$	37.9	$\boxed{41.7}$

It is interesting to note that the best value of r depends on the difference between the parameters.

IV
THE PARAMETERS OF THE HYPOTHETICAL DISTRIBUTION ARE ESTIMATED FROM THE SAMPLE

In the preceding the parameters of the hypothetical function were specified, which is a rather exceptional case in the applications.

We will now examine the case that the hypothetical distribution function including its shape parameter, if any, is specified, but that the location and scale parameters have to be estimated from the sample.

In order to emphasize that the class probabilities p_i are functions of the parameters μ and β, the definition (1) will

be written

$$X^2 = \Sigma[(v_i - N.p_i(\mu,\beta)^2/N.p_i(\mu,\beta)] \tag{35}$$

where

$$p_i = r_c - r_{c-1} = F[(x_c - \mu)/\beta] - F[(x_{c-1} - \mu)/\beta] \tag{36}$$

and x_c = the upper class limits

$$x_c = \beta.F^{-1}(r_c) + \mu \tag{37}$$

which specify the chosen division of the space of the variable x.

If the true values of μ and β are known, the value of X^2 is merely calculated by use of (35). In the present case, however, the parameters have to be replaced by their estimates $\hat{\mu}$ and $\hat{\beta}$.

Equ.(35) thus becomes

$$X^2 = \Sigma[(v_i - N.p_i(\hat{\mu},\hat{\beta}))^2/N.p_i(\hat{\mu},\hat{\beta})] \tag{38}$$

For a fixed set of class limits x_c the probabilites p_i will no longer be constant. If, however, the set r_c is fixed, then p_i will be constant, while the class limits x_c will vary from sample to sample. It is obvious that the sampling distribution of X^2 will depend upon the estimation method chosen.

The problem of finding the limiting distribution of X^2, when one or more of the parameters are estimated from the sample, was solved by R.A.Fisher [3] for a specific method of estimation, viz., the maximum likelihood method. He found that it is only necessary to reduce the number of degrees of freedom of the limiting χ^2-distribution by one unit for each parameter estimated from the sample. This simple and attractive rule is not valid, if other estimators are used.

Two alternative methods will be examined: the best linear and the pseudo-standardization methods.

4.1 Best linear estimation of the parameters

The distributions of $X^2_{r/N}$ are computed by use of Program 23/71. From each random sample of size N the parameters are estimated by use of the following formulas

$$\hat{\beta} = \Sigma c_i x_i$$
$$\hat{\mu} = \Sigma d_i x_i \quad (39)$$

The coefficients c_i, d_i are given by Sarhan & Greenberg [5] for the normal distributions and $N = 2(1)20$ and for the Weibull distributions by Weibull [6] for $N = 5(5)20$ and $\alpha = 0.05$, 0.1, 0.3, 0.5, 0.7, 1.0, 1.5, 2.0

Introducing these estimates into the formula

$$x_c = \hat{\beta} . F^{-1}(c/r) + \hat{\mu} \quad (c = 0,1,2,--,r) \quad (40)$$

the class limits and corresponding class frequencies v_i are obtained.

It is convenient to use instead of X^2 the statistic

$$K = N . X^2_{r/N}/2r = \Sigma v_i^2/2 - N^2/2r \quad (41)$$

because K is a non-negative integer and $\inf K = 0$

Some distributions

$$Q(x) = \mathrm{Prob}(K > x) \quad (42)$$

for the normal distribution are presented in Table 5.

A test of normality is easily performed by use of these tables. From the given sample under examination the value of K is computed and the probability of having a larger value is read from the table. If this probability is too small, the hypothesis of normality is rejected. The same test can be used for other hypothetical distributions, if the necessary tables have been prepared.

4.2 Pseudo-standardization of the samples

If the estimates $\hat{\mu}$ and $\hat{\beta}$ are replaced by x_1 and $(x_N - x_1)$, respectively, the pseudo-standardized variable

$$t_i = (x_i - x_1)/(x_N - x_1) \tag{43}$$

is obtained. Since $t_1 = 0$ and $t_N = 1$, the size of the transformed sample is equal to $(N - 2)$.

For properly chosen class limits t_c, the modified test statistic

$$K = \Sigma v_i^2/2 - (N-2)^2/2r \tag{44}$$

can be computed by use of Program 24/71.

In Table 6 the results for $N = 10$, $r = 2, 4, 8$ and for normal and exponential dbns are presented.

These tables can be used for testing normality and exponentiality.

4.3 The statistic $X^2_{r/N}$ as a shape operator

When $X^2_{r/N}$ and K are computed from samples which have location and scale invariance, as the two above-mentioned alternatives, both of them can be used as shape operators. For example, from the distributions in Table 6 it is easy to compute the power of deciding between normal and exponential distributions.

The result for $N = 10$ is

$$r = 2 \quad 4 \quad 8$$

$$DP = 35.5 \quad 41.9 \quad 39.0\%$$

This decision power is not very good, to some extent due to the fact that $X^2_{r/N}$ is independent of permutations in the observed values of v_i. For instance, for $r = 2$, $N = 10$, the set $v_1 = 0$; $v_2 = 10$ yields the same value of $X^2_{2/10}$ as $v_1 = 10$; $v_2 = 0$.

A substantial improvement can be obtained by separating the probabilities of

$$
\begin{array}{rlcl}
v_1, v_2 = & 0,\ 10 & \text{and} & 10,\ 0 \\
 & 1,\ 9 & & 9,\ 1 \\
 & 2,\ 8 & & 8,\ 2 \\
 & & \text{etc.} &
\end{array}
$$

By doing so the value above of DP = 35.5% was raised to DP = 55.0%. In some other cases, the improvement was still better. These results have motivated the introduction of a new test operator, denoted by VI. Its properties and usefulness will be demonstrated in a following Scientific Report.

Table I - The exact distribution $F(w)$ and the normal distribution $\phi(w)$ of the random variable $w = (v - v_o)/\sqrt{v_o}$ for $v_o = 1$ and various sample sizes N

v	w	F(w)							$\phi(w)$
		N = 5	10	20	50	100	1,000	10,000	-
0	-1	32.768	34.868	35.848	36.417	36.603	36.770	36.786	15.866
1	0	73.728	73.610	73.583	73.557	73.576	73.576	73.576	50.000
2	1	94.208	92.981	92.451	92.157	92.063	91.979	91.972	84.134
3	2	99.328	98.720	98.409	98.224	98.163	98.107	98.103	97.725
4	3	99.968	99.836	99.742	99.679	99.657	99.636	99.635	99.865
5	4	100.000	99.985	99.967	99.953	99.947	99.941	99.941	99.997
6	5	-	99.999	99.997	99.995	99.994	99.992	99.992	100.000
7	6	-	100.000	100.000	100.000	100.000	99.999	99.999	-
8	7	-	-	-	-	-	100.000	100.000	-
E(w) =		0	0	0	0	0	0	0	
E(w) =		0.80	0.90	0.95	0.98	0.99	0.999	0.9999	

Table II - The exact distribution $F(w)$ and the normal distribution $\phi(w)$ of the random variable $w = (v - v_o)/\sqrt{v_o}$ for $N = 1{,}000$ and $v_o = 1, 4$, and 16

$v_o = 1$

v	w	F(w)	$\phi(w)$
0	-1	36.770	15.866
1	0	73.576	50.000
2	1	91.979	84.134
3	2	98.107	97.725
4	3	99.636	99.865
5	4	99.941	99.997
6	5	99.992	100.000
7	6	99.990	-
8	7	100.000	-

$E(w) = 0$; $E(w^2) = 0.999$

$v_o = 4$

v	w	F(w)	$\phi(w)$
0	-2.0	1.817	2.275
1	-1.5	9.114	6.681
2	-1.0	23.752	15.866
3	-0.5	43.300	30.854
4	0.0	62.884	50.000
5	0.5	78.545	69.146
6	1.0	88.975	84.134
7	1.5	94.923	93.319
8	2.0	97.888	97.725
9	2.5	99.201	99.379
10	3.0	99.723	99.865
11	3.5	99.912	99.977
12	4.0	99.974	99.997
13	4.5	99.993	100.000
14	5.0	99.998	-
15	5.5	100.000	

$E(w) = 0$; $E(w^2) = 0.996$

Table II (Continued)

$v_o = 16$

v	w	F(w)	Φ(w)	v	w	F(w)	Φ(w)
0	- 4.00	0.000	0.003	19	0.75	81.393	77.336
1	- 3.75	0.000	0.009	20	1.00	86.996	84.134
2	- 3.50	0.001	0.023	21	1.25	91.248	89.434
3	- 3.25	0.008	0.058	22	1.50	94.324	93.319
4	- 3.00	0.037	0.135	23	1.75	96.451	95.993
5	- 2.75	0.130	0.298	24	2.00	97.859	97.725
6	- 2.50	0.380	0.621	25	2.25	98.753	98.777
7	- 2.25	0.957	1.223	26	2.50	99.298	99.379
8	- 2.00	2.122	2.275	27	2.75	99.618	99.702
9	- 1.75	4.210	4.007	28	3.00	99.799	99.865
10	- 1.50	7.575	6.681	29	3.25	99.897	99.942
11	- 1.25	12.499	10.566	30	3.50	99.949	99.977
12	- 1.00	19.098	15.866	31	3.75	99.975	99.991
13	- 0.75	27.253	22.664	32	4.00	99.988	99.997
14	- 0.50	36.601	30.854	33	4.25	99.994	99.999
15	- 0.25	46.593	40.130	34	4.50	99.997	100.000
16	0.00	56.595	50.000	35	4.75	99.999	-
17	0.25	66.009	59.870	36	5.00	100.000	-
18	0.50	74.368	69.146	-	-	-	-

$E(w) = 0; \; E(w^2) = 0.984$

Table III- Some exact distributions of the statistic $X^2_{r/N}$ - for completely specified hypothetical distributions

$X^2_{2/10}$

x	p(x)	P(x)	Q(x)
0.0	24.609	24.609	75.391
0.4	41.016	65.625	34.375
1.6	23.438	89.062	10.938
3.6	8.789	97.852	2.148
6.4	1.953	99.805	0.195
10.0	0.195	100.000	0.000

$E(X^2) = 1.0$; $Var(X^2) = 1.8$

$X^2_{2/20}$

x	p(x)	P(x)	Q(x)
0.0	17.620	17.620	82.380
0.2	32.036	49.656	50.344
0.8	24.027	73.682	26.318
1.8	14.786	88.468	11.532
3.2	7.393	95.861	4.139
5.0	2.957	98.818	1.182
7.2	0.924	99.742	0.258
9.8	0.217	99.960	0.040
12.8	0.036	99.996	0.004
16.2	0.004	100.000	0.000
20.0	0.000	100.000	0.000

$E(X^2) = 1.0$; $Var(X^2) = 1.9$

$p(x) = Prob(X^2 = x)$

$P(x) = Prob(X^2 \leq x)$

$Q(x) = Prob(X^2 > x)$

$X^2_{3/9}$

x	p(x)	P(x)	Q(x)
0.000	8.535	8.535	91.465
0.667	38.409	46.944	53.056
2.000	21.125	68.069	31.931
2.667	15.364	83.432	16.568
4.667	11.529	94.955	5.045
6.000	2.561	97.516	2.484
8.000	1.097	98.613	1.387
8.667	1.097	99.710	0.290
12.667	0.274	99.985	0.015
18.000	0.015	100.000	0.000

$E(X^2) = 2.0$; $Var(X^2) = 3.55556$

Table IV - Some M.C.-determined distributions of $X^2_{r/N}$ for completely specified hypothetical distributions

$X^2_{2/10}$

x	p(x)	P(x)	Q(x)
0.0	24.52	24.52	75.48
0.4	40.96	65.48	34.52
1.6	23.46	89.94	11.06
3.6	9.07	98.01	1.99
6.4	1.85	99.86	0.14
10.0	0.14	100.00	0.00
$E(X^2)$	= 0.99839		
r - 1	= 1.00000		
$Var(X^2)$	= 1.74553		
2(r-1)(N-1)/N	= 1.80000		

$X^2_{5/10}$

x	p(x)	P(x)	Q(x)
0	1.23	1.23	98.77
1	15.47	16.70	83.30
2	15.60	32.30	67.70
3	19.47	51.77	48.23
4	12.52	64.29	35.71
5	15.32	79.61	20.39
6	3.05	82.66	17.34
7	7.81	90.47	9.53
8	4.07	94.54	5.46
9	1.62	96.16	3.84
10	0.22	96.38	3.62
11	2.27	98.65	1.35
12	0.46	99.11	0.89
13	0.45	99.56	0.44
15	0.03	99.59	0.41
16	0.16	99.75	0.25
17	0.19	99.94	0.06
19	0.02	99.96	0.04
23	0.02	99.98	0.02
24	0.02	100.00	0.00
$E(X^2)$	= 3.98033		
r - 1	= 4.00000		
$Var(X^2)$	= 7.09714		
2(r-1)(N-1)/N	= 7.20000		

$p(x) = Prob(X^2 = x)$

$P(x) = Prob(X^2 \leq x)$

$Q(x) = Prob(X^2 > x)$

Table IV (Continued)

x	p(x)	P(x)	Q(x)
0	0.04	0.04	99.96
2	1.60	1.64	98.36
4	11.31	12.95	87.05
6	21.41	34.36	65.64
8	22.54	56.90	43.10
10	19.32	76.22	23.78
12	8.07	84.29	15.71
14	8.84	93.13	6.87
16	3.01	96.14	3.86
18	1.59	97.73	2.27
20	1.01	98.74	1.26
22	0.68	99.42	0.58
24	0.29	99.71	0.29
26	0.13	99.84	0.16
28	0.02	99.86	0.14
30	0.06	99.92	0.08
32	0.06	99.98	0.02
34	0.01	99.99	0.01
42	0.01	100.00	0.00

$E(X^2) = 8.98307$

$r-1 = 9.00000$

$Var(X^2) = 15.89416$

$2(r-1)(N-1)/N = 16.20000$

Table V - Some Q-functions of K for normal dbn and two best-linearly estimated parameters

K	N = 10			N = 20			
	r = 2	r = 5	r = 10	r = 2	r = 4	r = 10	r = 20
0	58.98	94.23	99.40	70.40	96.30	99.96	100.00
1	9.36	61.55	94.13	24.53	70.66	99.33	100.00
2	-	44.02	75.14	-	63.51	95.65	99.90
3	-	22.65	52.71	-	42.60	88.50	99.19
4	.47	15.20	28.73	5.25	36.44	78.00	96.53
5	-	7.12	13.49	-	27.37	65.34	89.92
6	-	6.09	7.93	-	24.70	52.82	78.41
7	-	2.59	3.04	-	14.17	41.94	63.47
8	-	.89	1.44	-	13.42	31.78	48.15
9	.00	.78	.76	.60	9.79	23.33	33.88
10	-	.73	.47	-	8.12	17.19	22.67
11	-	.27	.15	-	6.54	12.61	14.43
12	-	.16	.03	-	5.12	8.68	9.10
13	-	.06	.02	-	2.95	6.04	5.70
14	-	-	-	-	-	4.11	3.71
15	-	-	-	-	2.05	3.02	2.30
16	-	.05	.00	.01	1.96	2.20	1.30
17	-	.00	-	-	1.38	1.76	.83
18	-	-	-	-	1.10	1.24	.47
19	-	-	-	-	.56	.90	.31
20	-	-	-	-	.43	.58	.16
21	-	-	-	-	.27	.44	.09
22	-	-	-	-	.23	.37	.05
23	-	-	-	-	.16	.29	.03
24	-	-	-	-	-	.16	.02
25	-	-	-	.00	.07	.11	.01

$$K = \Sigma v_i^2/2 - N^2/2r$$

Table VI - Some Q-functions of K for pseudo-standardized sample from normal and exponential populations

K	Normal dbn			Exponential dbn		
	$r=2$	$r=4$	$r=8$	$r=2$	$r=4$	$r=8$
0	78.08	97.70	99.83	92.98	99.34	99.08
1	39.71	72.33	94.72	75.26	91.58	98.56
2	-	63.72	73.38	-	88.37	91.30
3	-	38.86	51.47	-	75.21	81.34
4	14.12	30.57	27.51	49.66	70.31	65.05
5	-	18.73	18.16	-	60.63	57.21
6	-	16.13	11.03	-	56.33	46.74
7	-	7.29	3.92	-	40.73	32.22
8	-	6.15	3.13	-	39.54	30.08
9	2.79	4.11	2.19	20.07	35.74	25.69
10	-	-	1.12	-	-	18.01
11	-	1.98	.30	-	23.50	11.50
12	-	.56	-	-	15.19	11.18
13	-	-	.21	-	-	10.19
15	-	-	.06	-	-	4.70
16	.00	-	.02	.00	-	3.02
17	-	.07	-	-	3.94	-
21	-	-	.00	-	-	.38
24	-	.00	-	-	.00	-
28	-	-	-	-	-	.00

$$K = \Sigma v_i^2/2 - (N-2)^2/2r$$

$r = 2$; $t_c = 0.50000$

$r = 4$; $t_c = 0.30500 \quad 0.50000 \quad 0.69500$

$r = 8$; $t_c = 0.18307, 0.30500, 0.40663, 0.50000$
$0.59337, 0.69500, 0.81693$

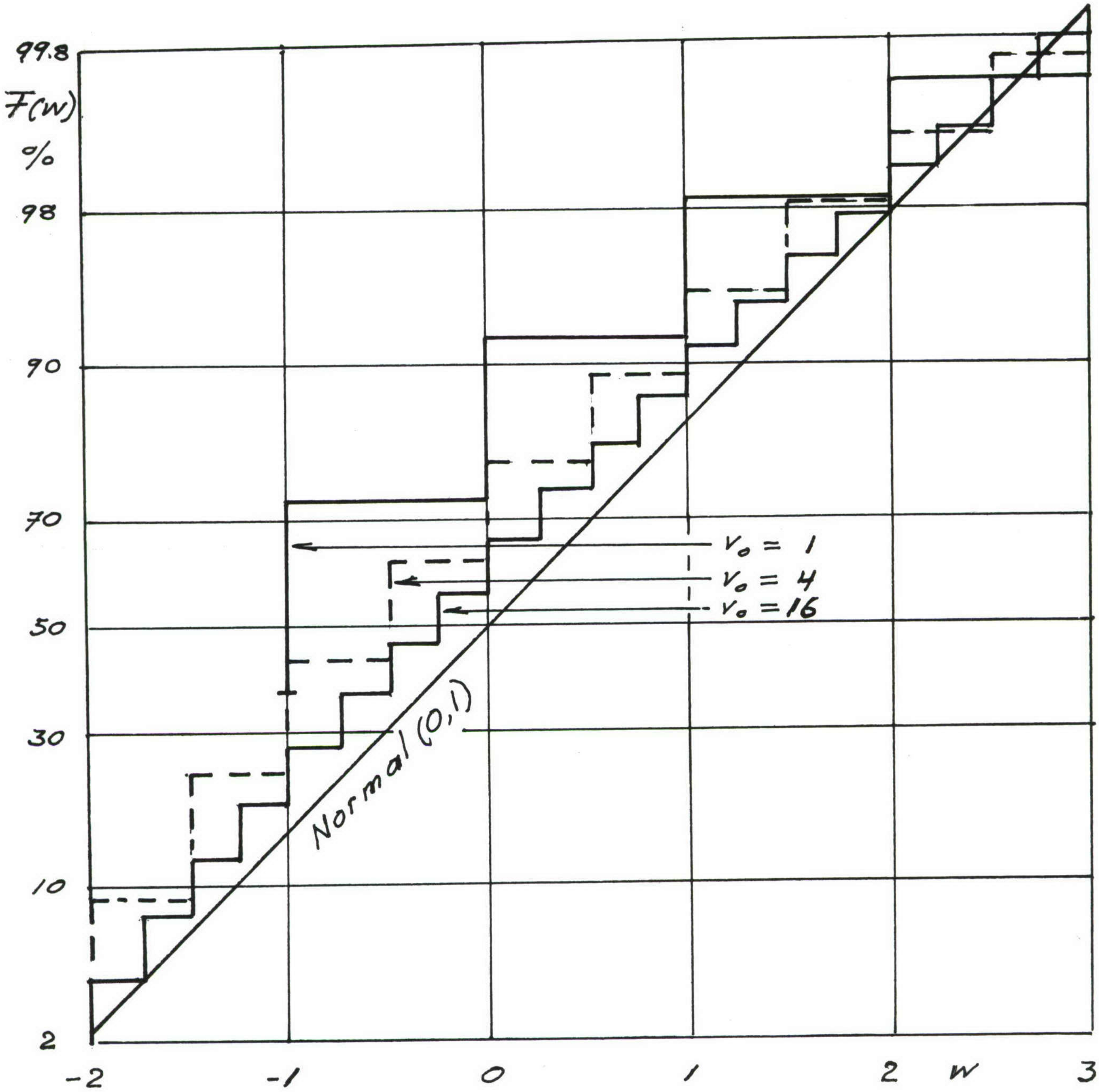

Fig. 1. Exact distributions $F(w)$; $N = 1{,}000$; $v_0 = 1, 4, 16$.

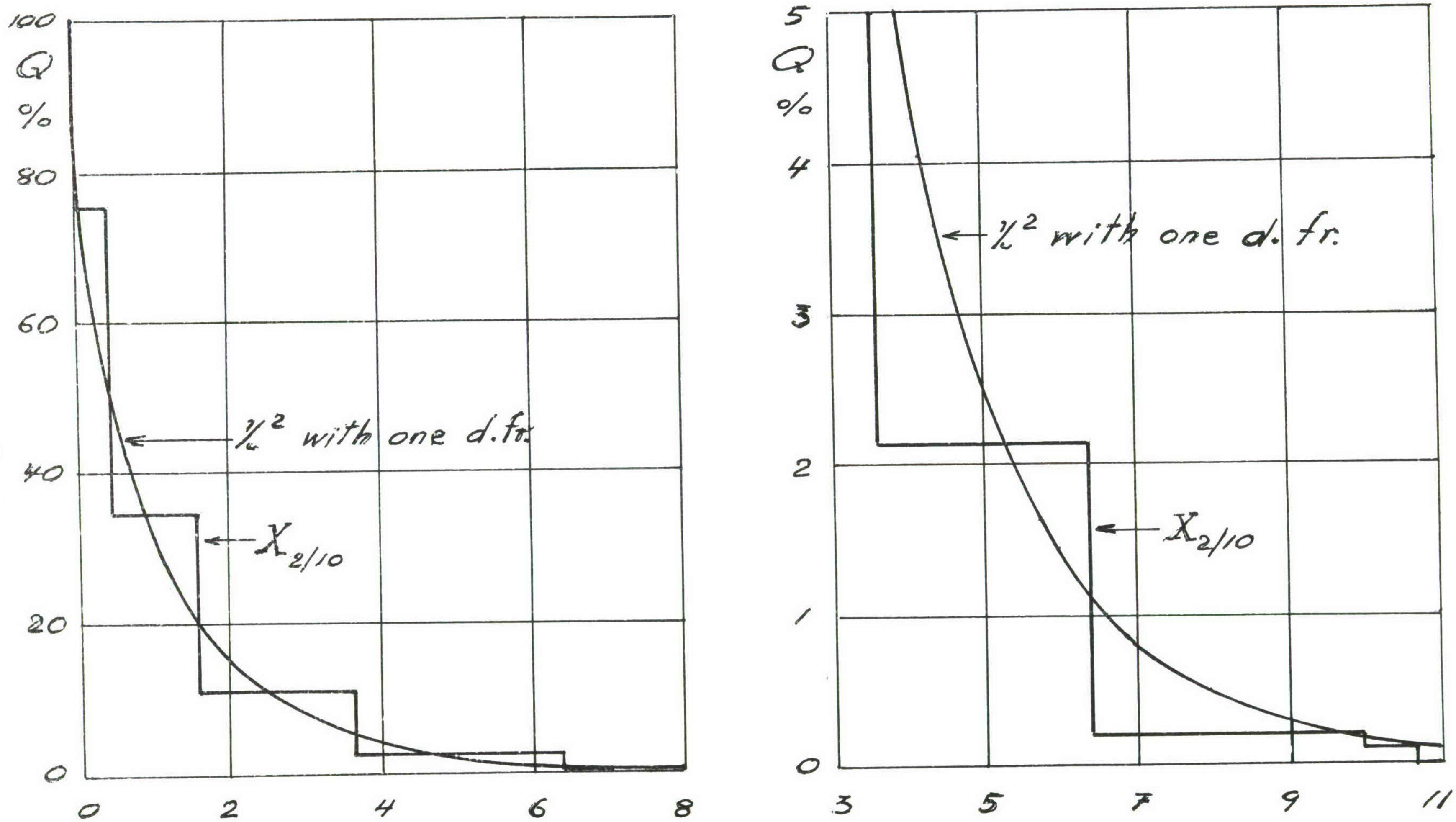

Fig. 2. Q-functions of $X^2_{2/10}$ and χ^2 with one degree of freedom.

REFERENCES

1.	Pearson, K.	On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Phil. Mag., V, 50, 1900, 137.
2.	Cramér, H.	Mathematical Methods of Statistics. Almqvist & Wiksell, Uppsala, 1945.
3a.	Fisher, R.A.	On the interpretation of χ^2 from contingency tables, and the calculation of P. J. Roy, Soc., 85, 1925, 87.
3b.	Fisher, R.A.	The conditions under which χ^2 measures the discrepancy between observations and hypothesis. J. Roy, Soc., 87, 1924, 442.
4.	Wilks, S.S.	Order statistics. Bull. Amer. Math. Soc. Vol. 54, No. 1, 6-50.
5.	Sarhan, A.E. & Greenberg, B.G.	Contributions to Order Statistics. Wiley & Sons, New York, 1962.
6.	Weibull, W.	Approximations to best linear, unbiased order statistic estimators. AFML-TR-67-198, 1967.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Professor Waloddi Weibull Ch. Fontanettaz 15 1012 Lausanne, Switzerland	UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

THE EKS-SQUARE TEST OF GOODNESS OF FIT - AN IMPROVEMENT OF THE CHI-SQUARE TEST

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Waloddi Weibull

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
May 1973	27	7

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
F 44620-72-C-0028	AFML-TR-73-94
b. PROJECT NO. 7351	
c. Task No. 735106	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Air Force Materials Laboratory(LL) Wright-Patterson AFB, Ohio 45433

13. ABSTRACT

When applying the classical Chi-square test of goodness of fit, it is always assumed that the test statistic is χ^2 - distributed. Since this is true only for very large samples, some restrictions on the class frequencies have to be introduced. It is generally accepted that none of the expected frequencies should be less than ten, which makes this test useless for small and moderate samples.

In order to eliminate these - from a practical viewpoint severe - restrictions, it is proposed to use the exact sampling distribution instead of the limiting χ^2 - distribution. When doing so, the test will be called the Eks-square test.

Programs have been written for computing these distributions and the improvements attained have been stated.

The possibilities of using the modified test statistic as a location, scale, and shape operator have been examined and illustrated by numerical examples. Several tables have been prepared.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Test Statistic						
Sampling Distribution						
Location, Scale, and Shape Operator						

☆U.S.Government Printing Office: 1973 — 759-494/582

UNCLASSIFIED

Security Classification